FINAL REPORT: Distribution approved for public release; distribution is unlimited

Bayesian Hierarchical Models to Augment the Mediterranean Forecast System

Ralph F. Milliff
Colorado Research Associates Division, NWRA
3380 Mitchell Lane
Boulder, CO 80301
phone: (303) 415-9701 fax: (303) 415-9702 email: *milliff@cora.nwra.com*

Christopher K. Wikle
Department of Statistics, University of Missouri
146 Middlebush Hall
Columbia, MO 65211
phone: (573) 882-9659 fax: (573) 884-5524 email: *wikle@stat.missouri.edu*

L. Mark Berliner
Department of Statistics, The Ohio State University
1958 Neil Ave.
Columbus, OH 43210
phone: (614) 292-0291 fax: (614) 292-2096 email: *mb@stat.osu.edu*

Emanuele Di Lorenzo
School of Earth and Atmospheric Sciences, Georgia Institute of Technology
311 Ferst Drive
Atlanta, GA, 30332
phone: (404) 894-3994 fax: (404) 894-5638 email: *edl@eas.gatech.edu*

Award Number: N00014-09-C-0485
http://www.cora.nwra.com/MedBhm/

17 February 2012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 03-02-2012	**2. REPORT TYPE** Final Report	**3. DATES COVERED** *(From - To)* 5/1/2009 – 5/1/2011

4. TITLE AND SUBTITLE Bayesian Hierarchical Models to Augment the Mediterranean Forecast System	**5a. CONTRACT NUMBER** N00014-09-C-0485
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Ralph F. Milliff	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NorthWest Research Associates PO Box 3027 Bellevue, WA 98009-3027	**8. PERFORMING ORGANIZATION REPORT NUMBER** NWRA-12-RB454
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research ONR 252, Phil Eisenhaur (CACI) 875 N. Randolph St., Room 1238W Arlington, VA 22203-1995	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The overall project goal has been to test the feasibility and practicality of Bayesian Hierarchical Model (BHM) methods in aspects of the Mediterranean Forecast System (MFS); an operational ocean data assimilation and forecast system.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	UL	1	**19b. TELEPHONE NUMBER** *(include area code)*

1 Introduction

The research to be summarized here was made possible by sustained funding from the Physical Oceanography Program of the U.S. Office of Naval Research (ONR), and through collaborations with scientists in the Operational Oceanography Group (GNOO; Grupo Nazionale di Oceanografia Operativa) and access to computing resources of the National Climate Center (CMCC: Centro euro-Mediterraneo per i Cambiamenti Climatici) of the Istituto Nazionale di Geofisica e Vulcanologia (INGV) in Bologna. This unique combination of resources has provided a platform from which we are launching multidisciplinary research efforts that are leading to broad-scale adoption of Bayesian Hierarchical Modeling (BHM) methods in oceanography and related fields. At the time of the initial funding, BHM methods were relatively unproven for applications in geophysical fluid settings with practical state- and data-space dimensions, and operational time constraints. The applications of BHM to be reported here demonstrate the practicality and advantages of the method for realistic problems in operational ocean forecasting.

Our research program goal has been to test the feasibility and practicality of BHM methods in aspects of the Mediterranean Forecast System (MFS); an operational ocean data assimilation and forecast system that produces 10-day forecasts for the state of the Mediterranean Sea every day. Three separate BHM developments address different aspects of operational ocean forecasting at INGV. In the MFS-Wind-BHM project, ensemble ocean forecast methods were developed based on posterior distributions of the surface vector wind (SVW) process over the Mediterranean Sea. In the MFS-Error-BHM project, time-dependent background error covariance information is provided to the sequential data assimilation system of MFS. Finally, multi-model and multi-parameter super-ensembles for target ocean processes have been the objective of the MFS-SuperEnsemble-BHM project.

Results for MFS-Wind-BHM are documented in companion papers (Milliff et al., 2011 and Pinardi et al., 2011) that have appeared in the *Quarterly Journal of the Royal Meteorological Society* (*QJRMS*). Research for MFS-Error-BHM and MFS-SuperEnsemble-BHM is ongoing as described below, with manuscripts to be submitted in calendar year 2012.

2 MFS-Wind-BHM

The companion papers, Milliff et al., (2011) and Pinardi et al., (2011) provide a full-scale demonstration of the practicality and advantages of BHM methods in operational ensemble ocean forecasting. Principal achievements and findings of these works include:

- A BHM for the SVW (MFS-Wind-BHM) uses multi-platform data stage inputs (ECMWF analyses and forecasts, and QuikSCAT SVW retrievals) to efficiently generate ensembles of vector winds, four-times daily, at $0.5°$ resolution for the entire Mediterranean Sea forecast domain. A snapshot of the SVW ensembles in the Western Mediterranean is shown in Figure 1.

- The SVW wind ensembles provide realistic estimates of the SVW (i.e. in the posterior mean sense) and SVW uncertainty (i.e. in the spread of the posterior distribution) given the

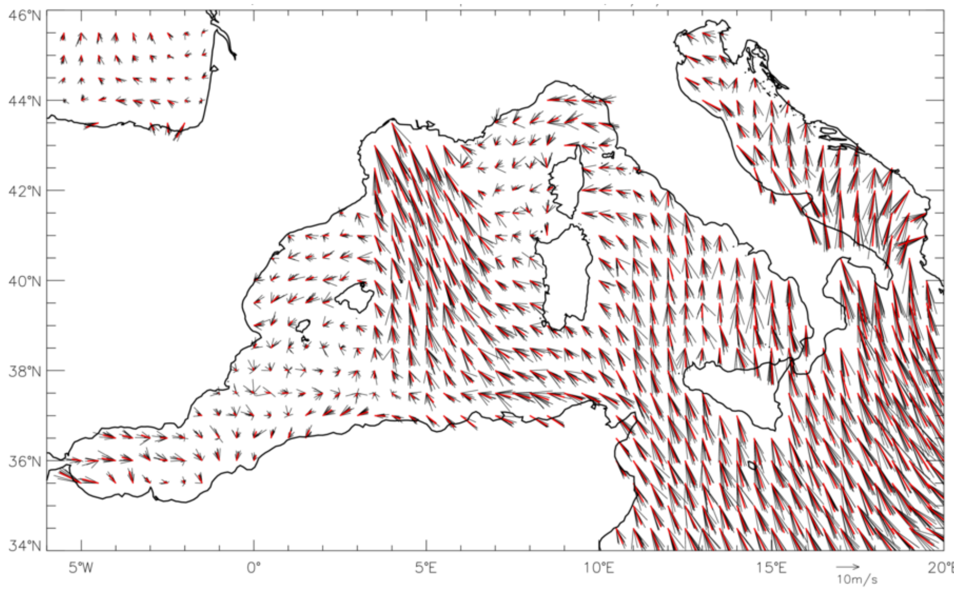


Figure 1: Sample SVW realizations from the posterior distribution of MFS-Wind-BHM for the western Mediterranean basin on 2 Feb 2005 at 1800 UTC. Ten wind vectors (black) are plotted at each grid location. A red vector at each location represents the posterior mean wind vector (see also Milliff et al., 2011).

data and the validity of the process model based on geostrophic and ageostrophic balances for the MFS domain taken as a whole, as functions of time.

- Realizations from the posterior distribution of MFS-Wind-BHM drive sequential data assimilation steps to generate ensemble ocean initial conditions that exhibit realistic and balanced spread in multivariate ocean fields including sea-surface temperature (SST), sea-surface height (SSH), ocean currents, etc. The ensemble initial condition spread is focussed on the scales of ocean mesoscale eddies. Initial condition spread in SST and SSH are depicted for 10-member ensembles in Figure 2.

- Realizations from the posterior distribution during the forecast period are based on data stage inputs from ECMWF forecasts (*vs.* analyses during the assimilation period). However, the ensemble spread from ocean forecasts continues to be concentrated on ocean mesoscales which, appropriately, are the most uncertain scales of the MFS forecasts.

- The MFS-Wind-BHM ocean ensemble forecast method is less arbitrary than random perturbation methods, and better at producing baroclinic perturbations (i.e. at the ocean pycnocline) in the ocean response than more traditional methods (e.g. as practiced at ECMWF). A comparison of spread in a density section (latitude *vs.* depth) during the forecast period is shown in Figure 3.

The companion papers (Milliff et al., 2011 and Pinardi et al., 2011) provide a practical example of *uncertainty quantification* via the BHM methodology for ocean-atmosphere systems of realistic scale. The implications of this demonstration extend to the climate system as well (e.g. see also Berliner and Kim, 2008).

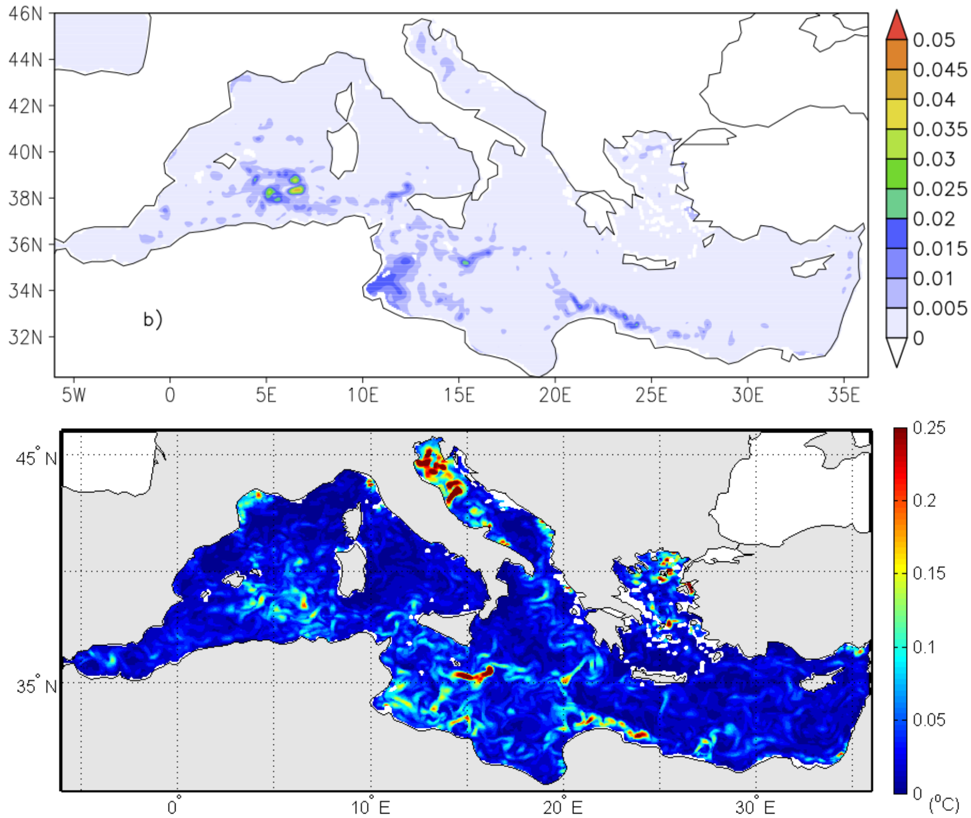


Figure 2: Standard deviations in ocean ensemble initial conditions for sea-surface height (top) and sea-surface temperature (bottom) for a 10-member ensemble forced during the data assimilation period (14 days) by realizations of the MFS-Wind-BHM posterior distribution.

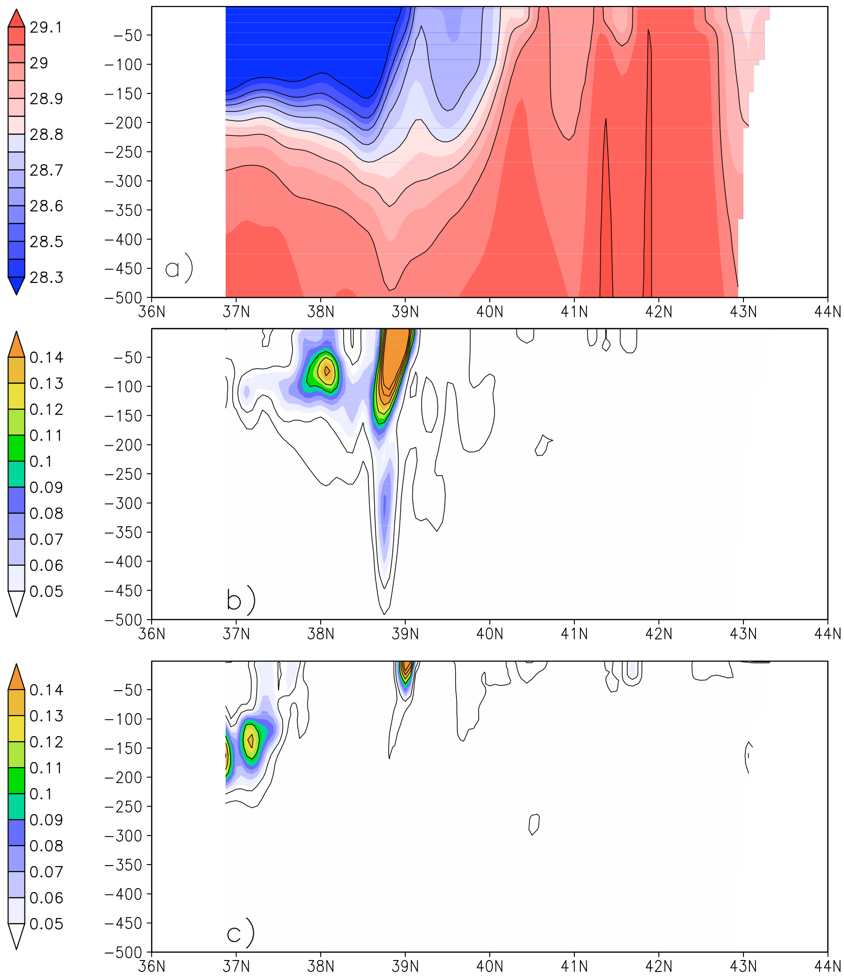


Figure 3: Meridional section of $\sigma = \rho - 1000\,(kg\,m^{-3})$ at $5^{\circ}E$ (Algerian coast to the left and French coast to the right) for 17 Feb 2005. Panel (a) is the daily mean σ for forecast day 10. The σ standard deviation for the ensemble forecast on day 10 as driven by realizations of the MFS-Wind-BHM posterior distribution is shown in panel (b), and the σ standard deviation forced by the ECMWF ensemble prediction system winds is shown in panel (c). The contour interval for σ standard deviations (b and c) is $0.01\,(kg\,m^{-3})$.

3 MFS-Error-BHM

The 3dVar MFS data assimilation system employs a multivariate background error covariance matrix $\mathbf{B}$, the vertical part of which (i.e. $\mathbf{B_v}$) weights model estimates of temperature (T) and salinity (S) profiles (Dobricic et al., 2005; 2007). In order to account for changes in regional water mass properties and seasonal variations that can be abrupt, MFS imposes *ad hoc* partitions of the Mediterranean Sea domain into 13 sub-regions and 4 seasons. A table of 13×4 $\mathbf{B_v}$ matrices is maintained and changes in $\mathbf{B_v}$ are imposed as step-functions from region to region, and from season to season.

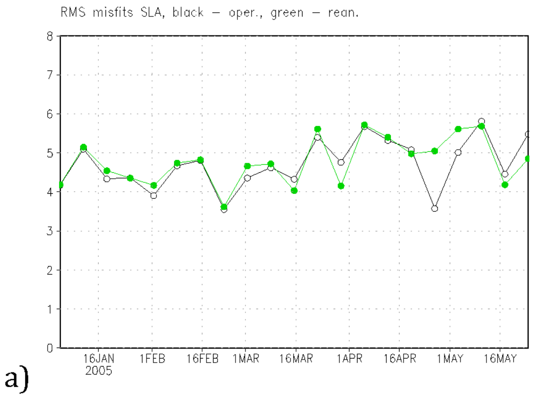
The purpose of MFS-Error-BHM is to develop a method for temporal variation in $\mathbf{B_v}(t)$ driven by data stage inputs from: a) forecast *vs.* data misfits, $\mathbf{d}$; and b) forecast anomalies $\mathbf{q}$, that are the year-day departures from forecast climatology for MFS. The misfits $\mathbf{d}$ mostly reflect forecast differences with respect to ARGO profiles at a few locations within each region during the data assimilation period. The ARGO data are sparse in space and time such that the $\mathbf{d}$ data are noisier than the climatology anomalies $\mathbf{q}$. MFS-Error-BHM is flexibly designed to weight $\mathbf{q}$ and $\mathbf{d}$ differently for each implementation of the model.

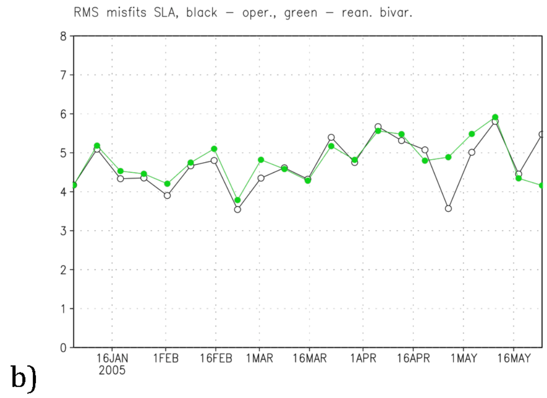
The vertical part of the forecast model error covariance is reduced in dimension by projecting onto vertical basis functions with time dependent amplitude coefficients. Time-dependence is modeled via an error process model for which the error covariance is given by $\mathbf{B_v}(t)$. Details are provided in Wikle et al., (2012).

The time-dependent $\mathbf{B_v}(t)$ from MFS-Error-BHM is compared against the operational system in (retrospective) reforecast experiments spanning several seasons. Metrics for comparison include: time-histories of region-average RMS differences in sea-level anomaly (SLA) with respect to analyzed satellite data; and time- and region-averaged vertical profiles of RMS misfits in T and S. A growing matrix of developmental reforecast runs have been performed in the Gulf of Lyon region of the MFS domain (i.e. region 3). In addition to testing developments in MFS-Error-BHM, these experiments have served to refine the $\mathbf{d}$ and $\mathbf{q}$ datasets. Reforecasts with $\mathbf{B_v}(t)$ based on vertical structure functions computed from region-average $T(z)$ and $S(z)$ profiles have not shown marked improvement over the operational system (i.e. with fixed seasonal $\mathbf{B_v}$) at MFS. In retrospect, we note that the target scale for the error covariance matrix in the MFS 3dVar is the ocean mesoscale. In computing vertical structure functions that are the basis of MFS-Error-BHM from region-average profiles, we have washed out important variability signals that are focused at the ocean mesoscale (i.e. day to day eddy field variability).

We are in the process now of recomputing vertical structure functions based on the $T(z)$ and $S(z)$ profiles at each grid location (i.e. $5140\, x, y$) in region 3. Variability associated with mesoscale eddies will be reflected in the vertical structures derived from this larger reanalysis dataset. In addition, we are incorporating sea-surface height (SSH) analyses at each grid location into the covariance matrix structure (i.e. adding another row and column to $\mathbf{B_v}$). SSH provides an vertically-integrated signal of the $T(z)$ and $S(z)$ variations in the upper ocean.

In arguing for the mesoscale enhancements of the vertical structure functions in MFS-Error-BHM, Dr. Srdjan Dobricic (INGV lead) performed some sensitivity tests with fixed $\mathbf{B_v}$ in reforecast experiments with the MFS operational system. Figure 4 documents the impact of mesoscale





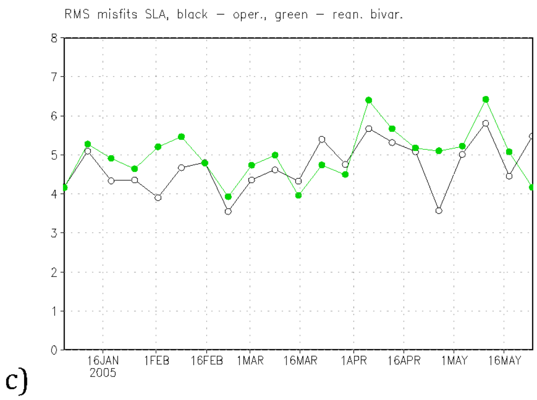


Figure 4. Sea level anomaly root mean square differences for 3 reforecast experiments, using 3 versions of the time-dependent error covariance matrix from MFS-Error-BHM: a) tri-variate, mesoscale; b) bi-variate, mesoscale; c) bi-variate, spatially averaged (see text for details).

variability in $\mathbf{B}$ in reforecast experiments for the period January-May 2005[1]. Three panels plot the root-mean-square (RMS) difference region-average sea-level anomaly (SLA) comparing forecasts with different vertical structure functions used in computing $\mathbf{B}$ (green traces) *versus* RMS SLA difference for operational MFS (black traces). Panel (a) in Fig. 4 compares the RMS SLA traces for a version of $\mathbf{B_v}$ wherein vertical structure functions are computed from $T(z)$, $S(z)$ and SSH at every grid location (i.e. 5140 locations) within the region; thus preserving and emphasizing the ocean mesoscale variability. The RMS SLA is comparable to, but not yet better than the operational RMS SLA. However, time dependence has not yet been tested here via MFS-Error-BHM. Panel (b) in Fig. 4 depicts the RMS SLA comparison for a version of $\mathbf{B_v}$ for which SSH is not considered in computing the vertical structure functions. The comparison with operational RMS SLA is slightly degraded from the comparison in panel (a). Finally, the version of $\mathbf{B_v}(t)$ tested in panel (c) is based on vertical structure functions computed from regional average $T(z)$ and $S(z)$ as is the case in MFS-Error-BHM. This washes out important variability associated with the ocean mesoscale and the test case (green) is worse than the operational case (black).

[1]It has recently been discovered that there is a mismatch in the year of the reforecast experiment shown here and the year for which vertical basis functions were computed for $\mathbf{B_v}$. These experiments are in the process of being re-run now.

4 MFS-SuperEnsemble-BHM

The Berliner and Kim (2008) BHM has been reformulated to address target ocean processes on daily and sub-seasonal timescales as they are simulated in operational and experimental forecast models at MFS; i.e. the Ocean PArallelise (OPA; Madec et al., 1998), and Nucleus for European Modeling of the Ocean (NEMO; Madec, 2008) models, respectively. The reformulation is being implemented in a proof-of-concept calculation using daily temperature and salinity profiles (i.e. $T(z,t)$ and $S(z,t)$) for a location in the Rhodes Gyre region of the Eastern Mediterranean Sea, during February and March 2006. These months span the period within which Levantine Intermediate Water (LIW) typically forms in the Rhodes Gyre. Colder and saltier intrustions in $T(z,t)$ and $S(z,t)$, between the surface and about $400\,m$, are the signals of LIW formation and spreading in the Rhodes Gyre region.

Simulations from OPA and NEMO provide data stage inputs for the multi-model ensemble state estimation BHM. Ensembles are generated following the methodology in Milliff et al., 2011. Ten realizations of a posterior distribution for the surface wind are used to generate 10 members each, of the 11 member ensembles for OPA and NEMO. The eleventh member for each ensemble is forced by ECMWF winds that were used to spin up each model to the 1 February 2006 start date for the experiment. Simulation results are collected for the MFS grid location at $26.875°E$, $33.5°N$. Figure 5 depicts every other member of the OPA and NEMO ensembles, in $T(z,t)$ and $S(z,t)$ that serve as data stage inputs to the BHM.

Details of the model implementation, including full-conditional distribution specifications, will be provided in Berliner et al., 2012. A brief description of the BHM design is as follows. We let the form of the data stage distribution be:

$$\mathbf{Y}_{t,m,i_m}|\mathbf{B}_{t,m}, \mathbf{X}_t, \mathbf{\Sigma}_{Y_m} \sim N_d(\mathbf{X}_t + \mathbf{B}_{t,m},\, \mathbf{\Sigma}_{Y_m}) \tag{1}$$

where there are $m = 1, \ldots, M$ models (e.g. $M = 2$ for OPA and NEMO), and $i_m = 1, \ldots, R$ replicates for each model (i.e. $R = 11$ for 10 replicates driven be winds from realizations from MFS-Wind-BHM, and the 11th replicate driven by ECMWF winds).

Let the target ocean process vector be $\mathbf{X}_1, \ldots, \mathbf{X}_T$, where each instance of $\mathbf{X}_t$ is d-dimensional, say for $d/2$ depths, and the period $T = 60d$. So, in our proof-of-concept calculation, $\mathbf{X}$ will be the distributions for $T(z,t)$ and $S(z,t)$, at the point of interest in the Rhodes Gyre. The process model is given by a first-order multivariate autoregression (AR-1):

$$\mathbf{X}_t - \boldsymbol{\theta}_t = \mathbf{H}(\mathbf{X}_{t-1} - \boldsymbol{\theta}_{t-1}) + \boldsymbol{\epsilon}_t, \quad t = 2, \ldots, T \tag{2}$$

where a prior mean vector $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_T$ is removed at each time step, $\mathbf{H}$ is the autoregression transition matrix, and the $\boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \mathbf{\Sigma}_\epsilon)$ are the innovation vectors. Here, we use the Sys3a version of the MFS ocean analyses for $\boldsymbol{\theta}_t$. Figure 6 depicts the Sys3a analysis $T(z,t)$ and $S(z,t)$ at the point of interest in the Rhodes Gyre.

At the next level of the BHM hierarchy, the distribution for the $\mathbf{B}_{t,m}$ in (1) are specified. $\mathbf{B}_{t,m}$ are used to account for inherent smoothing in time, and offsets in amplitudes, that are characteristic of data stage inputs from each forecast model. These are the so-called model bias

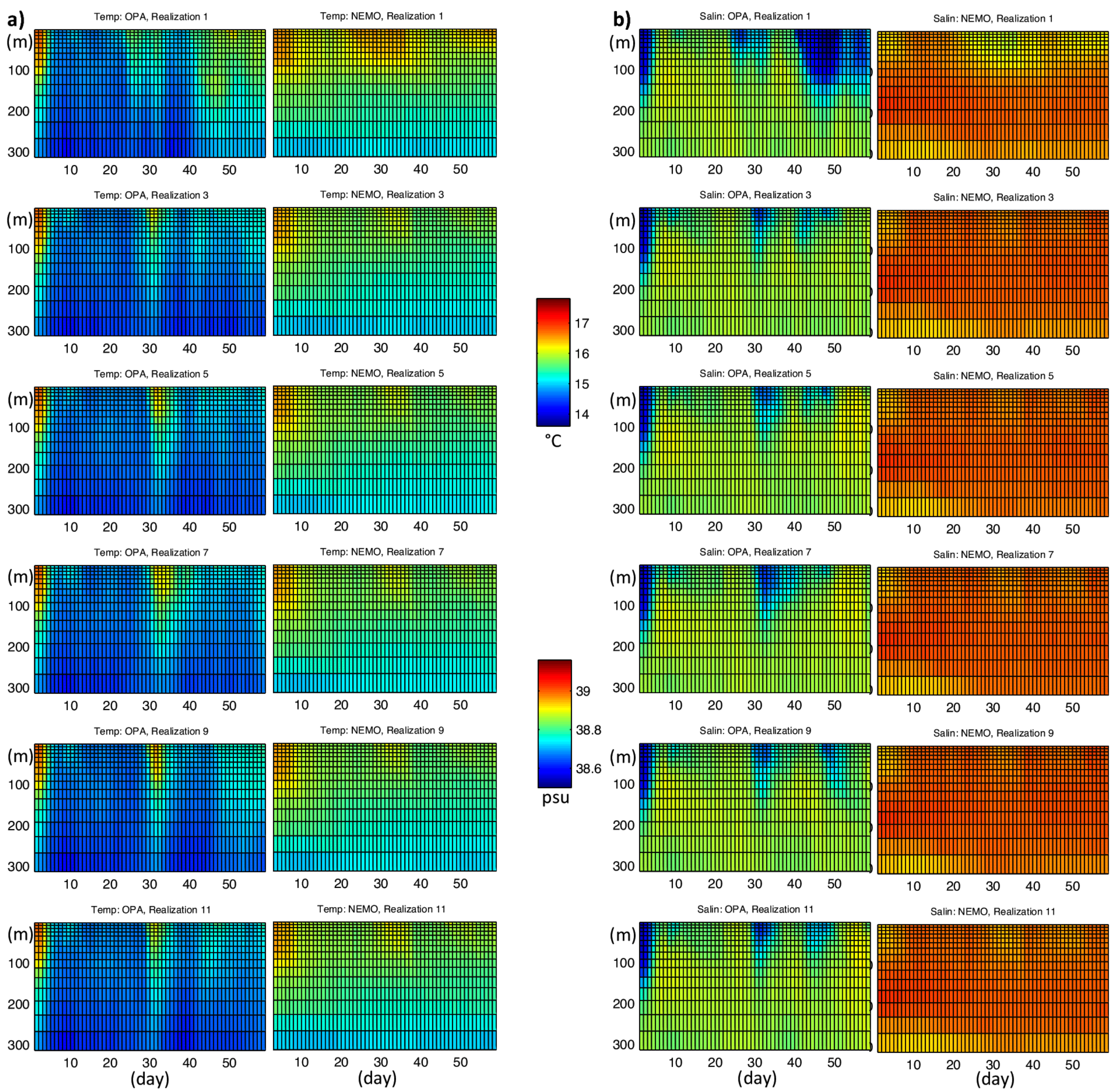


Figure 5: Every other ensemble member for a) $T(z,t)$ (left column OPA and right column NEMO) and b) $S(z,t)$ (left OPA, right NEMO). These fields enter the data stage model of the multi-model ensemble state estimate BHM.

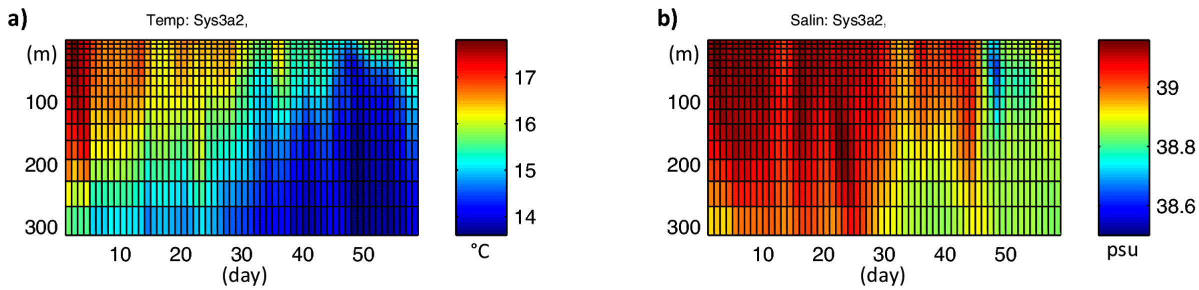


Figure 6: Sys3a analysis a) temperature and b) salinity from the INGV operational system. The Sys3a analysis $T(z,t)$ and $S(z,t)$ are used as prior means to demonstrate the multi-model ensemble state estimation BHM methodology.

parameters that will be estimated for each model in the posterior distribution.

Berliner et al. (2012) theorize that either a complete set of relevant observations or a strong prior are required to reduce uncertainty and identify model biases in the posterior distribution. Here, we use an observation-based prior mean, $\boldsymbol{\theta}_t$ in (2), to demonstrate the method. Table 1 depicts the layout of the random variables in the BHM (from Berliner et al., 2012).

Table 1: Layout of all Variables

State vector	$\mathbf{X}$	$\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_T$
Model class mean	$\boldsymbol{\beta}$	$\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_T$
Model 1	Ensemble	$\mathbf{Y}_{111}, \mathbf{Y}_{211}, \ldots, \mathbf{Y}_{T11}$ $\mathbf{Y}_{112}, \mathbf{Y}_{212}, \ldots, \mathbf{Y}_{T12}$ $\vdots$ $\mathbf{Y}_{11n_1}, \mathbf{Y}_{21n_1}, \ldots, \mathbf{Y}_{T1n_1}$
	Offset	$\mathbf{B}_{11}, \mathbf{B}_{21}, \ldots, \mathbf{B}_{T1}$
	Piece-wise constant	$\mathbf{b}_{11}, \ldots, \mathbf{b}_{s_1 1}$
	Prior Mean Offset	$\boldsymbol{\mu}_{11}, \ldots, \boldsymbol{\mu}_{s_1 1}$
	$\vdots$ $\vdots$	$\vdots$ $\vdots$
Model M	Ensemble	$\mathbf{Y}_{1M1}, \mathbf{Y}_{2M1}, \ldots, \mathbf{Y}_{TM1}$ $\vdots$ $\mathbf{Y}_{1Mn_M}, \mathbf{Y}_{2Mn_M}, \ldots, \mathbf{Y}_{TMn_M}$
	Offset	$\mathbf{B}_{1M}, \mathbf{B}_{2M}, \ldots, \mathbf{B}_{TM}$
	Piece-wise constant	$\mathbf{b}_{1M}, \ldots, \mathbf{b}_{s_M M}$
	Prior Mean Offset	$\boldsymbol{\mu}_{1M}, \ldots, \boldsymbol{\mu}_{s_M M}$

The posterior distribution for $T(z,t)$ and $S(z,t)$ are summarized in Figure 7. Posterior mean temperature and salinity profile evolutions are very similar to the Sys3a analyses. This is consistent with the differences between the OPA and NEMO ensembles, and their differences with

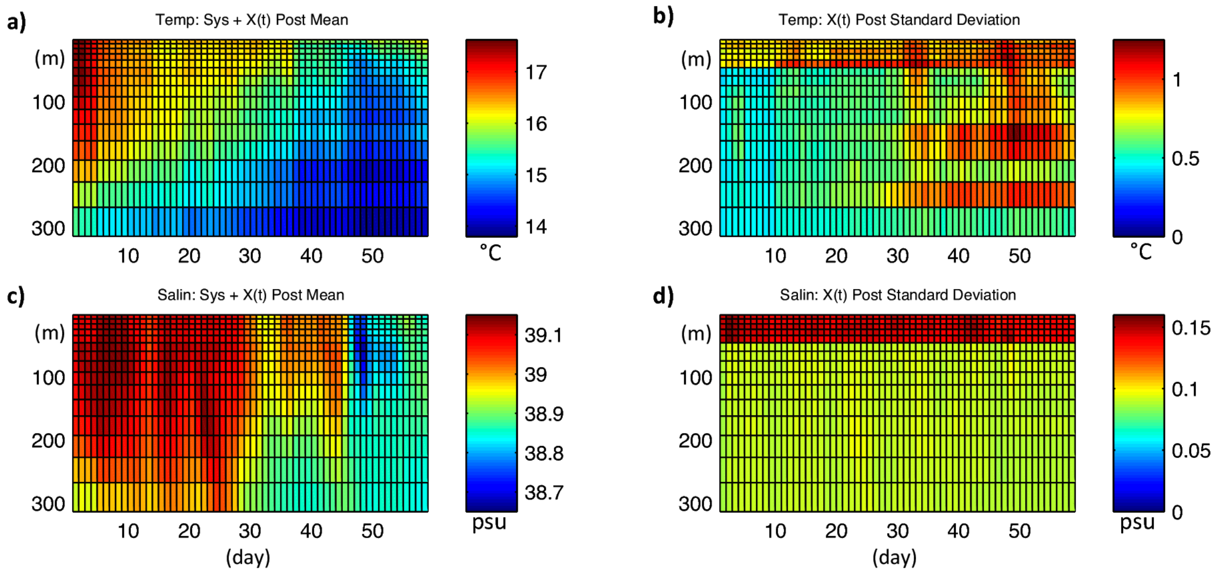


Figure 7: Posterior mean a) temperature and b) temperature uncertainty, c) salinity and d) salinity uncertainty from the multi-model ensemble estimate BHM.

respect to the prior mean from Sys3a. The time- and depth-dependent uncertainties reflect times and depths where the OPA and NEMO ensembles were most variable.

Figure 8 depicts summaries of the information from the BHM posterior distribution regarding OPA and NEMO biases, and the uncertainties in those biases, for both $T(z,t)$ and $S(z,t)$. In temperature, the NEMO model is too warm at depth, late in the February-March period. Conversely, the OPA model is too cold from the surface to about $200\,m$ in the early part of the period. Moreover, the uncertainty in the OPA temperature bias is greatest during the early period as well.

The practical interest in the proof-of-concept developments for the multi-model ensemble state estimation BHM have to do with probabilistic hindcasts of T and S signatures of LIW formation in the Rhodes Gyre. This is an important ocean process that, once the ocean is preconditioned, can occur in intermittent and sudden episodes, in response to extreme atmospheric forcing events. The SVW realizations from MFS-Wind-BHM are used to force replicates for each model in the multi-model superensemble during February and March (i.e. after pre-conditioning). The multi-model ensemble state estimation BHM bounds the uncertainty in the critical forcing events, and provides estimates of model biases in the simulation of important ocean processes. Clearly, a very similar model framework can be used to bound uncertainty in different, user-specified, ocean target processes (e.g. thermocline position and strength, transport across a predefined line or position, etc.), for ensembles from a wide variety of forward models.

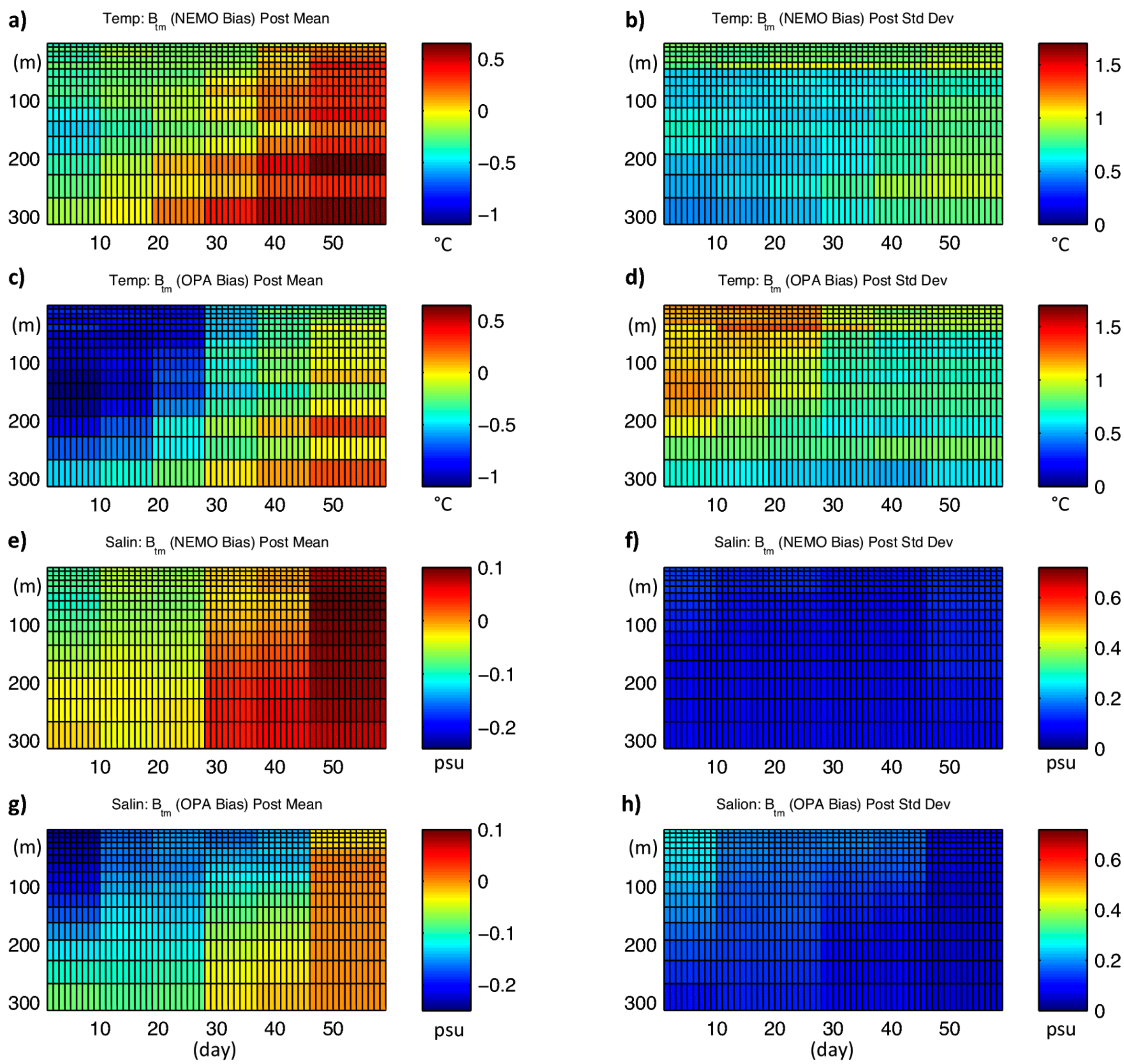


Figure 8: Posterior mean bias fields, and bias field uncertainties for: (a-b) temperature in the NEMO model; (c-d) temperature in the OPA model; (e-f) salinity in NEMO; and (g-h) salinity in OPA.

5 Med-ROMS Developments

The results described in the previous section demonstrate an ocean state estimation BHM given multi-model simulations and data. A superensemble ocean forecast BHM is a direct extension of that work. Toward that end, in the final two years of funding, the project supported the development of a new ocean forecast system for the Mediterranean Sea based on the Regional Ocean Modeling System (ROMS). Med-ROMS (*www.med-roms.org*) has an average horizontal resolution of $8.6\,km$ with 30 terrain following layers. The western boundary of the model includes a region of open ocean where a nudging open ocean boundary is used to introduce fluxes of momentum and buoyancy associated with the exchange with Atlantic waters. A recent report of the ROMS model numerics and configurable options is given Haidvogel et al. (2008) or at the ROMS official website (*myroms.org*).

Med-ROMS data archive period 2000-2006

MED-roms Forecasting
http://www.MED-roms.org

RUN NAME	WINDS	HEAT FLUX	FRESHWATER FLUX	IC
SPINUP	NCEP Climatology	NCEP Climatology + NOAA SST correction Climatology	NCEP Climatology + SODA SSS correction Climatology	SODA Climatology
QSC	QSCAT (daily 2000-2005)	NCEP (monthly 2000-2005) + NOAA SST correction (monthly 2000-2005)	NCEP (monthly 2000-2005) + SODA SSS correction Climatology	Feb. 1, 0006 SPINUP
QSC-*IC*	QSCAT (daily 2000-2005)	NCEP (monthly 2000-2005) + NOAA SST correction (monthly 2000-2005)	NCEP (monthly 2000-2005) + SODA SSS correction Climatology	Jan. 1, 2000 MFS-SYS2B Analysis
EWF-*IC*	ECMWF (daily 2001-2005)	NCEP (monthly 2000-2005) + NOAA SST correction (monthly 2000-2005)	NCEP (monthly 2000-2005) + SODA SSS correction Climatology	Jan. 15, 2001 MFS-SYS2B Analysis
EWF-*sys2b*	ECMWF (daily 2001-2005)	MFS-SYS2B Analysis (monthly 2001-2005) + SST correction (monthly 2001-2005)	NCEP (monthly 2001-2005) + MFS-SYS2B Analysis SSS correction (monthly 2001-2005)	Jan. 15, 2001 MFS-SYS2B Analysis
QSC-*sys2b*	QSCAT (daily 2000-2005)	MFS-SYS2B Analysis (monthly 2001-2005) + SST correction (monthly 2001-2005)	NCEP (monthly 2001-2005) + MFS-SYS2B Analysis SSS correction (monthly 2001-2005)	Jan. 15, 2001 MFS-SYS2B Analysis

To build robust statistics for LIW and its model representation error we have generated an extensive data archive by driving Med-ROMS with different surface and boundary fluxes of momentum and buoyancy. These are summarized in Table 2 and include different surface momentum fluxes such as the reanalysis from ECMWF and National Center for Environmental Prediction (NCEP), as well as the QuikSCAT satellite derived winds and the MFS Sys2b analysis. The MFS analysis and the Simple Ocean Data Assimilation (SODA) products were also used to prescribe the fluxes at the western open boundary of Med-ROMS.

As an example, Figure 9 shows and compares the horizontal spread of the $250\,m$ salinity in the March mean for the different Med-ROMS simulations (see Table 2 for details on the boundary conditions), and for the Med-ATLAS observations (*www.ifremer.fr/medar*) and MFS operational analyses. As suggested earlier, $250\,m$ salinity is a good proxy for LIW and reveals some of the biases in the exchange dynamics between the eastern and western Mediterranean basins. Given the strong and prolonged changes in LIW over the period 2000-2006, we also performed three longer term 40-year hindcasts of the Mediterranean circulation by forcing Med-ROMS with ECMWF fluxes. These long-term integrations provide information on the natural range of temporal variability of LIW and have also served to diagnose important forcing mechanisms of decadal scale variations of the circulation. Specifically, we found that although most of the interannual variability is controlled by the influence North Atlantic Oscillation, some of the long-term large-scale changes in the Mediterranean circulation are remotely driven by decadal-El Niño variations in the tropical Pacific. These latter results are beyond the initial scope of the project but are a direct consequence of the Med-ROMS development. A detailed report of these findings and of Med-ROMS performance is available in Di Lorenzo et al. (2012).

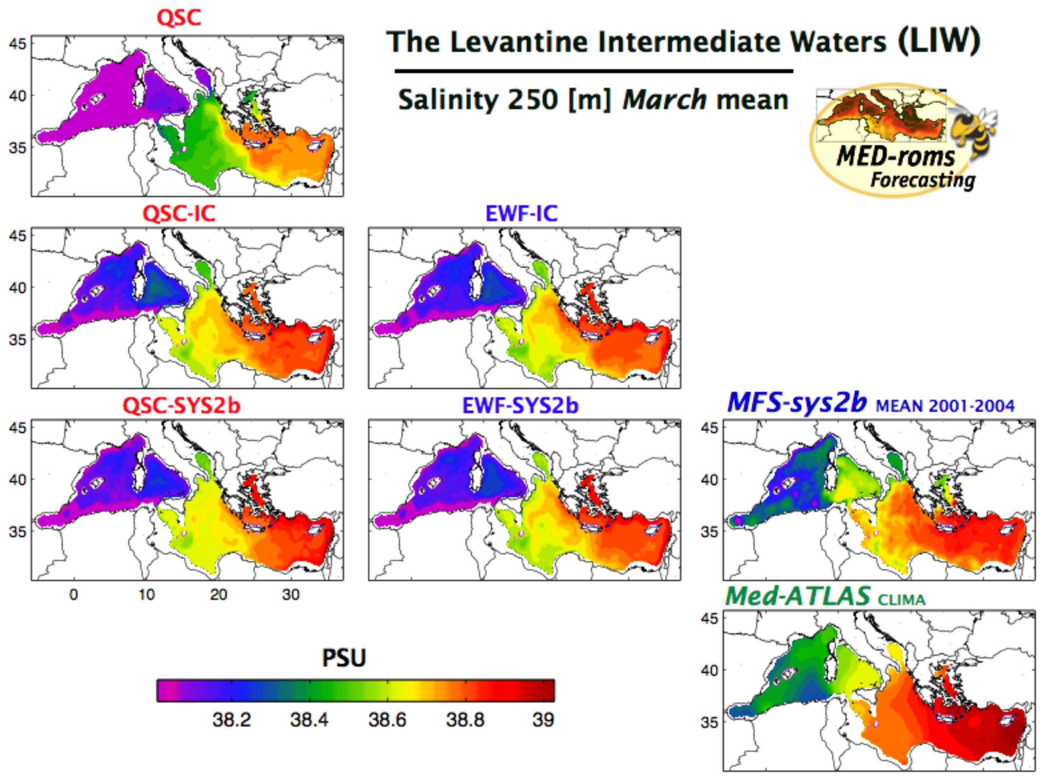


Figure 9: LIW in Med-ROMS inferred from salinity at 250 m from simulations that use different boundary conditions (QSC, EWF). These are compared to the Med-ATLAS observations and the MFS operational analyses.

Med-ROMS was also used to perform an ensemble of 11 simulations for the period of 2006. These data archives along with the other Med-ROMS long-term integrations have been made publicly available on the Georgia Tech OpenDAP server (*data.eas.gatech.edu/med.php*).

6 Extending BHM in Realistic Settings

Given the initial support from ONR, scientific applications of BHM in large state-space systems have expanded to include projects supported by several agencies, covering a wide variety of topics. These include:

- *Ocean ecosystem parameter estimation:* US Globec funding from NSF.
- *Forecasting ocean ecosystem indicators with climate-driven process models:* Workshop funding from US Globec (DiLorenzo).
- *Bayesian hierarchical climate prediction:* funding from NSF (Wikle, Berliner).
- *Characterizing uncertainty in the impact of global change on large river fisheries; Missouri River sturgeon example:* funding from USGS (Wikle).

- *A BHM for the Madden-Julian Oscillation:* funding from NASA International Ocean Vector Winds Science Team (IOVWST).
- *A global surface wind BHM:* funding from NASA IOVWST.
- *A regional ocean surface flux BHM (for the Mediterranean):* funding from NASA IOVWST.
- *Characterizing irreducible model error in ocean forecast systems:* funding from ONR Basic Research Challenge.

The surface flux BHM noted above continues our collaboration with INGV, further extending the work supported by ONR in the projects reported here.

7 References

Berliner, L.M., C.K. Wikle, R.F. Milliff and N. Pinardi, 2012: "Bayesian multi-model ensembling for ocean state estimation", *in prep.*

Berliner, L.M. and Y. Kim, 2008: "Bayesian design and analysis for superensemble-based climate forecasting", *J. Climate*, **21**, 1891-1910.

DiLorenzo, E.A., N. Pinardi and R.F. Milliff, 2012: Decadal forcing of the Mediterranean surface and subsurface circulation, *in prep*.

Dobricic, S., N. Pinardi, M. Adani, A. Bonazzi, C. Fratianni and M. Tonani, 2005: "Mediterranean Forecasting System: A new assimilation scheme for sea-level anomaly and its validation", *Quart. J. Royal Met. Soc.*, **131**, 3627-3642.

Dobricic, S., N. Pinardi, M. Adani, C. Fratianni, A. Bonazzi and V. Fernandez, 2007: "Daily oceanographic analyses by the Mediterranean Forecasting System at basin scale", *Ocean Sci.*, **3**, 149-157.

Haidvogel, D.B., and 19 co-authors, 2008: Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System, *J. Comp. Physics*, **227**, 3595-3624.

Milliff, R.F., A. Bonazzi, C.K. Wikle, N. Pinardi and L.M. Berliner, 2011: "Ensemble ocean forecasting, Part I: Ensemble Mediterranean winds from a Bayesian hierarchical model", *Quart. J. Royal Met. Soc.*, **137**, 858-878.

Pinardi, N., A. Bonazzi, S. Dobricic, R.F. Milliff, C.K. Wikle and L.M. Berliner, 2011: "Ensemble ocean forecasting, Part II: Mediterranean Forecasting System response", *Quart. J. Royal Met. Soc.*, **137**, 879-893.

Wikle, C.K., S. Dobricic, R.F. Milliff, N. Pinardi and L.M. Berliner, 2012: Hierarchical Bayesian specification of time-varying covariances for data assimilation, *in prep.*

Project Publications

Berliner, L.M., C.K. Wikle, R.F. Milliff and N. Pinardi, 2012: "Bayesian multi-model ensembling for ocean state estimation", *in prep.*

Berliner, L.M. and Y. Kim, 2008: "Bayesian design and analysis for superensemble-based climate forecasting", *J. Climate*, **21**, 1891-1910.

Berliner, L.M. and C.K. Wikle, 2007: Approximate importance sampling Monte Carlo for data assimilation, *Physica D*, **230**, 37-49.

Cressie, N. and C.K. Wikle, 2011: **Statistics for Spatio-Temporal Data**, John Wiley & Sons.

DiLorenzo, E.A., N. Pinardi and R.F. Milliff, 2012: Decadal forcing of the Mediterranean surface and subsurface circulation, *in prep*.

Gladish, D.W., C.K. Wikle and S.L. Holan, 2012: Covariate-based parameterization of time-varying spatial error covariances, *in prep*.

Holan, S. and C.K. Wikle, 2010: Comments on: "Bayesian Source Detection and Parameter Estimation of a Plume Model Based on Sensor Network Measurements" by C. Huang et al., *Applied Stochastic Models in Business and Industry*, **26**, 353-357.

Malmberg, A., A. Arellano, D.P. Edwards, N. Flyer, D. Nychka and C.K. Wikle, 2008: Interpolating fields of carbon monoxide data using a hybrid statistical-physical model, *The Annals of Applied Statistics*, **2**, 1231-1248.

Milliff, R.F., A. Bonazzi, C.K. Wikle, N. Pinardi and L.M. Berliner, 2011: Ensemble ocean forecasting, Part I: Ensemble Mediterranean winds from a Bayesian hierarchical model, *Quart. J. Royal Met. Soc.*, **137**, 858-878.

Pinardi, N., A. Bonazzi, S. Dobricic, R.F. Milliff, C.K. Wikle and L.M. Berliner, 2011: Ensemble ocean forecasting, Part II: Mediterranean Forecasting System response, *Quart. J. Royal Met. Soc.*, **137**, 879-893.

Wikle, C.K., S. Dobricic, R.F. Milliff, N. Pinardi and L.M. Berliner, 2012: Hierarchical Bayesian specification of time-varying covariances for data assimilation, *in prep.*

Wikle, C.K. and S.H. Holan, 2011: Polynomial nonlinear spatio-temporal integro-difference equation models, *Journal of Time Series Analysis*, **32**, 339-350; DOI: 10.1111/j.1467-9892.2011.00729.x.

Wikle, C.K. 2010: Low rank representations for spatial processes, In: *Handbook of Spatial Statistics*, A.Gelfand, P. Diggle, M. Fuentes, P. Guttorp (eds), Chapman and Hall. 107-118.

Wikle, C.K. and M.B. Hooten, 2010: A general science-based framework for spatio-temporal dynamical models, Invited discussion paper for: *Test*, **19**, 417-451.

Wikle, C.K. and R. F. Milliff, 2009: Invited discussion of "A spatio-temporal model for mean, anomaly and trend fields of North Atlantic sea surface temperature" by R.T. Lemos and B. Sanso, *Journal of the American Statistical Association*, **104**, 22.

Wikle, C.K. and L. M. Berliner, 2007: A Bayesian tutorial for data assimilation, *Physica D*, **230**, 1-16.

Project Presentations and Outreach

2005	
Jul	Confab in Boulder[1]
Oct	Milliff 2 week visit to INGV; seminars at INGV, U. Bologna, Ravenna
2006	
Jan	Milliff seminar U. Hawaii
Mar	Milliff seminar U. Washington, Dynamics Seminar series
May	Milliff seminars at UC Berkeley, UCSC
May	Milliff and Pinardi seminar at ONR HQ in DC
Jul	Milliff (Intro to BHM) seminar at NASA, OVWST meeting, Salt Lake City
Jul	Confab in Boulder
2007	
Jan	A. Bonazzi (U. Bologna) begins 12 month visit to NWRA/CoRA and UCSC
Apr	Bonazzi and Milliff presentations, European Geophysical Union Meeting, Vienna
May	Wikle invited seminar, SIAM Meetings, Snowbird
May	Wikle invited seminar, Harvard Univ. Climate Lecture series, Cambridge
Aug	Confab in Boulder
Aug	Milliff presentation, AMS Meeting, Portland, OR
Sep	Milliff seminar, Atmos. Sci. Dept., U. Washington
Nov	Milliff seminar, NRL Stennis (host: Gregg Jacobs)
2008	
Mar	Bonazzi thesis filed at U. Bologna
Mar	Milliff public outreach/invited talk, Boulder Torch Club
May	Milliff presentation, ONR Review, Scripps Inst Oceanog, La Jolla, CA
Aug	Wikle and Berliner invited talk, Joint Statistical Meetings, Denver
Aug	Confab in Boulder
Oct	Wikle public outreach/invited talk, Truman State Univ., Kirksville, MO
Nov	Milliff presentation (MFS-Wind-BHM), OVWST meeting, Seattle, WA
Dec	Wikle and Milliff invited presentations, NRC Workshop on Uncertainty Management in Remote Sensing of Climate Data, Wash. D.C.
2009	
Feb	DiLorenzo visit to NWRA/CoRA to setup MedROMS
Feb	Wikle public outreach/invited talk, Hickman High School, Columbia, MO
Apr	Milliff visits INGV, seminar
Apr	Wikle, invited seminar, Dept. Statistics, Univ. Wyoming
Jun	Milliff presentation, ONR Review in Chicago, IL

Aug	Wikle, invited seminar, Joint Statistics Meetings, Wash. D.C.
Aug	Wikle, JASA invited discussioin, Joint Statistics Meetings, Wash. D.C.
Aug	Confab in Boulder
Sep	Wikle, invited seminar, SAMSI Pgm on Space-Time Analysis (opening workshop)
Oct	Wikle, invited seminar, Iowa State Univ.

2010

Feb	Fiadeiro, Milliff, Wikle, session organizers; Berliner, Pinardi presenters
	Probabilistic Models in Ocean Science session, Ocean Sciences Meeting, Portland, OR
May	Wikle presentation, Inst. for Pure and Applied Math., Univ. California, Los Angeles
Aug	Confab in Boulder
Sep	Wikle, invited seminar, Dept. Statistics, Kansas State Univ.

2011

Apr	Wikle, invited seminar Dept. Biostatistics, Univ. Iowa
Aug	Confab in Boulder
Sep	Wikle, invited seminar, Norwegian Computing Center, Univ. Oslo
Sep	Ocean DA Workshop, UMd (Gregg Jacobs, Organizer), Milliff session chair
Oct	Wikle, invited seminar, Courant Inst. Applied Math., New York Univ.
Oct	Wikle, invited seminar, Dept. Applied Math., Univ. California, Santa Cruz
Nov	Wikle, invited seminar, Dept. Statistics, Univ. Illinois, Urbana-Champaign
Nov	Milliff presentation, ONR Review, Denver, CO

[1] Annual Confabs include all PI's, ONR representatives and invited guests. The discussions are lively and broad-ranging. Presentations of research issues and unsolved problems are encouraged over finished talks. Confab guests of relevance to the ONR BHM to Augment the MFS funding include INGV scientists (Pinardi, Dobricic, Oddo, Bonazzi), NRL scientists (Jacobs, Coelho, Richman), ONR-funded scientists (Moore, Powell). The Confabs have been expanded to include researchers from the projects spawned by initial BHM work described in this report (see section 6).